THE LANCET Oncology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Chang JY, Senan S, Paul MA, et al. Stereotactic ablative radiotherapy versus lobectomy for operable stage I non-small-cell lung cancer: a pooled analysis of two randomised trials. *Lancet Oncol* 2015; published online May 14. http://dx.doi.org/10.1016/S1470-2045(15)70168-3.

Supplemental methods

Surgery and SABR procedures

Surgery

Both open thoracotomy and video assisted thoracotomy (VATS) were acceptable procedures. Surgery may consist of a lobectomy, sleeve resection, bilobectomy or pneumonectomy as determined by the attending surgeon based on the operative findings. The type of resection chosen should provide complete removal of the primary lesion with negative gross and microscopic margins. Documentation of margins (bronchial and vascular and any other required) by frozen sections at surgery was recommended. Limited resection including wedge or segmental resections should not be performed unless there are unforeseen problems at the time of surgery. All accessible hilar (level 10) lymph nodes must be dissected from the specimen by the surgeon and submitted to the pathologist. A complete mediastinal lymph node sampling should be performed for each patient. For right-sided lesions, this includes 4R, 7, and 9. For left-sided lesions, this includes 5, 6, 7, and 9.

Post-operative adjuvant chemotherapy was permitted if there was no lymph node involvement in the final pathological stage. Post-operation radiotherapy and chemotherapy can be considered for a positive margin or pathological N1 and N2.

SABR

STARS Cyberknife based SABR

All patients on this protocol were treated with the Cyberkinfe radiosurgery system. The Cyberknife system uses projection X-ray for tumor localization and tracking. Patients in this protocol had 1–5 gold fiducials implanted inside or in the vicinity of tumor. Patients who were eligible for the fiducial-less "x sight" tracking (peripheral lesion with tumor size > 1.5 cm that can be directly visualized by projection X-ray) did not require fiducials. A CT simulation study was required for treatment planning with a 4-dimensional (4-D CT) being strongly recommended to verify that the fiducial system moved with the tumor and to study the relationship between the tumor and normal anatomy as a function of breathing phase. In the case of "X sight" based setup and tracking, the 4-D CT allowed visible tumor position verification throughout its range of motion. A CT image acquired without IV contrast was required for treatment planning dose calculation. It was recommended that a second CT image with IV contrast should be acquired and registered to the primary image for normal tissue and tumor delineation

There were 4 dose regimes allowed in the STARS trial. The choice of dose regime was dependent of the calculation algorithm and the tumor location. Two locations types were specified central and peripheral. Peripheral tumors were those located more than 2 cm away, using the CT lung window/level, in all directions around the proximal bronchial tree (carina, right and left main bronchi, right and left upper lobe bronchi, intermedius bronchus, right middle lobe bronchus, lingular bronchus, right and left lower lobe bronchi), major vessels, esophagus, heart, tracheal, vertebral body, pericardium, mediastinal pleural and brachial plexus. Central lesions were those that were not peripheral.

Two dose calculation algorithms were available on the Cyberknife treatment planning system, Monte Carlo which is a newer algorithm that properly accounts for the loss of scatter in the lungs, and Pencil Beam, the older and more widely available algorithm that does not account for the loss of scatter and thus overestimates the dose to small tumors in the lung. Both of these algorithms use CT-based attenuation corrections. Patients whose dose was calculated with Monte Carlo had prescriptions doses of 54 Gy in 18 Gy/fraction for peripheral lesions and 50 Gy in 12.5 Gy fractions for central lesions. Patients whose dose was calculated with Pencil Beam had prescriptions doses of 60 Gy in 20 Gy/fraction for peripheral lesions and 60 Gy in 15 Gy fractions for central lesions. Only institutions with Monte Carlo dose calculation were permitted to enroll patients with central lesion < 2 cm in diameter. All fractions were delivered on consecutive days.

For both peripheral and central lesions, the prescription isodose lines was required to cover 100% of the GTV covered and more than 95% of the PTV1 (GTV + 3 mm). In addition, PTV2, GTV + CTV margin (8 mm edited as necessary to account for physical boundaries) + 3 mm PTV margin, should be evaluated to make sure that more than 95% PTV2 is covered by at least the 50 Gy isodose lines regardless of calculation algorithm. Higher isodoses (hotspots) were manipulated to occur within the target and not in adjacent normal tissue. The maximal dose point must was kept to be within 30% of prescribed dose. All critical organ dose-volume limits were evaluated and respected.

ROSEL SABR

All patients on the ROSEL protocol where treated with linac based SABR. Patients who were treated with gating or tracking had radiopaque markers implanted in or near the tumor. All patients underwent CT simulation with 4D-CT scanning being strongly recommended in order to account for tumor motion. Acquisition of a slow-CT scan, or multiple rapid planning scans covering the entire tumor motion was allowed. The application of measures to reduce internal organ motion was optional with each system requiring approval of the study organizers.

As in the STARS trial there were 4 dose regimes allowed in the ROSEL trial. The choice of the dose regimes was dependent on the calculation algorithm and the tumor location. Conservative fractionation was used for tumors in contact with the chest wall or when the PTV extended to less than 5 mm from the mediastinal pleura. Standard fractionation was used for all other tumor locations. In contrast to the STARS study only lesions located ≥ 2 cm distal to the hilar structures on the diagnostic CT scan were eligible to inclusion.

As in the STARS trial the dose fractionation schedules allowed within the ROSEL trial were dependent on the dose calculation algorithm used. Pencil beam like algorithms (referred to as Type A algorithms in the ROSEL trial) had prescription doses for standard fractionation of 60 Gy at 20 Gy/fraction and for conservative fractionation of 60 Gy at 12 Gy/fraction. Monte Carlo and Convolution superposition algorithms (Type B) had prescription doses for standard fractionation of 54 Gy at 18 Gy/fraction and for conservative fractionation doses of 60 Gy at 12 Gy/fraction. The inter-fraction interval was kept to a minimum of 40 hours and a maximum of 4 days. The standard fractionation was delivered over 5-8 days, while the conservative fractionation was delivered over 10-14 days. Plan was generated such that the prescription isodose lines covered 95% of the PTV. In addition 99% of the PTV received at least of 90% of the prescription dose. The preferred maximum dose within the PTV was not <110% or >140% of the prescribed dose.

For both studies, adjuvant Chemotherapy after SABR was not permitted for stage I patients. For a patient who had persistent or recurrent local disease after SABR as demonstrated by CT and/or PET and confirmed by biopsy after SABR, salvage surgical resection was considered. Post-operation radiotherapy and chemotherapy were considered for a positive margin or pathological N1 and N2.

Data collection

Merge provided a database in which to enter all STARS study data. This database provided electronic case report forms (e-CRFs), which were accessible via a password-protected internet portal. All patient information complied with the Health Insurance Portability and Accountability Act. ROSEL study data were collected on paper forms (CRFs), entered into separate databases by two different operators, and, after comparison to prevent data entry errors, were merged with the STARS database.

Follow-up and treatment evaluation

Local recurrence was defined as CT evidence of progressive soft-tissue abnormalities in the treated lobe over time that corresponded to avid areas on PET/CT images (maximum standardized uptake values [SUV_{max}] >5) at >6 months after SABR. Biopsy was strongly recommended in such cases to confirm suspected recurrence. Recurrence in untreated lobes was scored as distant metastasis. Regional failure was defined as relapse in any intrathoracic lymph node outside of the PTV.

Objectives and statistical considerations

STARS is a 2-armed randomized Phase III trial in which randomization is to either stereotactic ablative radiotherapy (R) or surgical resection (S). The primary endpoint is overall survival (OS). Evaluation of Non-Target Lesions Marginal Failure (MF) refers to a measurable tumor located in the same lobe after SRT and within 2.0 cm of the treated planning target volume (PTV) that has been confirmed with PET with uptake of a similar intensity as the pre-treatment staging PET, or confirmed by biopsy. Regional Failure (RF) refers to a measurable tumor within lymph nodes along the natural lymphatic drainage typical for the location of the treated primary disease with dimension of at least >1.0 cm in the short axis on imaging studies (preferably CT scans) within the lung, bronchial hilum, or the mediastinum. Equivocally appearing enlarged lymph nodes should be positive on PET imaging or biopsied to confirm involvement with carcinoma. Metastatic Dissemination (MD) refers to the appearance after protocol therapy of cancer deposits characteristic of metastatic dissemination from non-small cell lung cancer. Appropriate evaluations for making this determination include physical examination and imaging studies. PET scan or biopsy to confirm MD is encouraged but not required.
The goal is to compare OS in the two arms, with particular attention to showing that the OS for arm R is not worse than the OS for arm S. Denote by πS and πR the 3-year OS in arms S and R, respectively. We will test the hypotheses:
H0: $\pi S \geq \pi R + 0.10$
H1: $\pi S < \pi R + 0$ .10.
Although the hypotheses are stated in terms of 3-year OS the analysis will be based on times to event. Assuming exponential OS and 3-year OS of 0.82 for arm S the non-inferiority margin of 0.10 at 3 years corresponds to a hazard ratio ($\lambda R/\lambda S$) of 1.66. That is, we are testing the hypothesis that the hazard ratio of arm R to arm S is less than 1.66. The hypotheses can be stated equivalently as
H0: $\lambda R/\lambda S \geq 1.66$
H1: $\lambda R/\lambda S < 1.66$.
All deaths are counted as events in the primary OS endpoint, regardless of cause. No specific adjustments/analyses will be done for competing risks. The DSMB was presented with cause of
death tables by arm, and similarly for eventual publications. Patients who are randomized to surgery (S) but who refuse this treatment and choose CyberKnife (R) were treated with CyberKnife (R) as described in the protocol. All patients are included in the analyses as randomized, however, following the intention-to-treat principle. Initially we expect approximately 85% of patients to be compliant with their randomized treatment, but we expect the compliance rate to increase as the study progresses.
The anticipated patient enrolment is 420 patients. At study conception of the STARS, the expected accrual rate was 60 patients per year. However, the study failed to meet this accrual goal. The dataset for the STARS trial included 36 patients and 24 variables. Originally 16 patients were randomized to the surgery arm and 20 patients to the SABR arm; however, one patient refused surgery after the randomization and received SABR instead. Following the principle of intention to treat, we included this patient in the surgery group. Another patient was randomized to surgery, but lobectomy was aborted when a metastatic mediastinal lymph node was identified; again, following the principle of intention to treat, we included this patient in the surgery group in the analysis.
ROSEL Power analyses: The study was set up to test whether radiosurgery is not inferior to the current standard (surgery), with the inferiority limit considered acceptable being a hazard ratio of 1.35. The 2-year local control for the surgery alone arm was estimated at 80%. A hazard ratio of 1.35 for the radiosurgery arm would result in a 74% 2 year local control, which was considered acceptable as surgical salvage is possible in patients who develop a local or hilar recurrence, particularly as radiosurgery has been shown to not result in significant deterioration in pulmonary function tests. When the sample size in each group would be 480, with a total number of events required of 200, a test of non-inferiority of the survival curve for the radiosurgery group to the survival curve for the surgery group with a 0.10 one-sided significance level would have 80% power to reject the null hypothesis of inferiority (a hazard ratio of 1.35 or greater) when the true hazard ratio is 1.0; this assumed an accrual period of 24 months and a maximum follow-up time of 36 months.

ROSEL Time schedule: An estimated 375 patients per year with stage IA disease are available for inclusion in the Netherlands. The study was planned to accrue patients from 2008 to 2010. Full follow up and data analysis for the primary end-points was planned to be completed by the end of 2012.

The SAS 9.1.3 (SAS, Cary, NC) and S-Plus 8.0 (TIBCO Software Inc., Palo Alto, CA) statistical analysis packages were used for the analyses.

Supplemental results

Patient characteristics and overall survival analysis between STARS and ROSEL

We conducted OS analysis for each trial separately. We found that for the STARS trial, the estimated OS rates at 1 year and 3 years were 100% (95% confidence interval [CI] 1,1) and 100% (95% CI 1,1) in the SABR arm and 81% (95% CI 0.64,1) and 67% (95% CI 0.47,0.96) in the surgical arm. The difference in OS between the two groups was statistically significant (log-rank P=0.0067). For the ROSET trial, the estimated OS rates at 1 year and 3 years were 100% (95% confidence interval [CI] 1,1) and 89% (95% CI 0.71,1) in the SABR arm and 100% (95% CI 1,1) and 100% (95% CI 1,1) in the surgical arm. The difference in OS between the two groups was not statistically significant (log-rank P=0.78).

Appendix 1:

Patient characteristic by study (ROSEL vs. STARS)

Variable	levels	Trial = ROSEL	Trial = STARS	p-value
gender	F	7(31.8%)	18(50%)	.1749
	M	15(68.2%)	18(50%)	.
Histology	Adenocar	3(37.5%)	26(72.2%)	.0378
	Other	2(25%)	1(2.8%)	.
	Squamous	3(37.5%)	9(25%)	.
Tumor Site	LLL	7(31.8%)	4(11.1%)	.4317
	LUL	5(22.7%)	10(27.8%)	.
	RLL	2(9.1%)	4(11.1%)	.
	RML	1(4.5%)	4(11.1%)	.
	RUL	7(31.8%)	14(38.9%)	.
Peripheral	No	.(.%)	5(13.9%)	.1455
	Yes	22(100%)	31(86.1%)	.
Pathological T	T0	1(9.1%)	.(.%)	.0535*
	T1	9(81.8%)	7(50%)	.
	T2	1(9.1%)	7(50%)	.
Treatment arm	SABR	11(50%)	20(55.6%)	.6807
	Surgery	11(50%)	16(44.4%)	.

*Only patients randomized to the Surgery arm had this information. Fisher's exact test or the Chi-square test was used to evaluate the association between study and other patient characteristics.

Appendix 2.

Wilcoxon rank sum test was used to evaluate the difference in the distributions of age and tumor size between the two trials.

Variable	dataset	n	mean	std	min	q1	median	q3	max	P value
age	ROSEL	22	66.7061	6.259	52.8624	62.3025	65.473	72.1971	75.9097	0.6364
	STARS	36	67.6242	9.9791	43.3018	62.5763	67.4839	73.5168	84.7721	.
Tumor_size_pathology_arm	ROSEL	11	1.8091	0.7739	0	1.5	2	2.1	3	0.1031
	STARS	14	2.3857	0.7492	1.1	1.8	2.5	3	3.5	.

Appendix 3.

Site principal investigators

Site (STARS)	Site Principal Investigator
Advocate Christ Medical Center, IL	Paul J. Gordon, M.D. (surgeon)
Central Baptist Hospital, Lexington, KY	Alan C. Beckman, M.D. (radiation oncologist)
Anova Cancer Care, Denver, CO	Gregg A. Dickerson, M.D. (radiation oncologist/CyberKnife)
Ruikang Hospital of Guangxi Traditional Chinese Medical University, China	Jian Liang, M.D. (surgeon)
University of Texas MD Anderson Cancer Center, Houston, TX	Jack A. Roth, M.D. (surgeon)
Centre Antoine Lacassagne, Nice, France	Pierre-Yves Bondiau, M.D., Ph.D. (SABR)
Oscar Lambret Cancer Center, Lille, France	Eric Lartigau, M.D., Ph.D. (radiation oncology)
Parkview Comprehensive Cancer Center, Fort Wayne, IN	Brian K. Chang, M.D. (CyberKnife)
Penrose Cancer Center, Colorado Springs, CO	Alan Monroe, M.D. (radiation oncologist)
St. Joseph Mercy Hospital, Ann Arbor, MI	Walter M. Sahijdak, M.D. (randiation oncologist)
St. Louis University, St. Louis, MO	Bruce J. Walz, M.D. (radiation oncologist)
Baylor-St. Luke's Medical Center, Houston, TX	Larry S. Carpenter, M.D. (radiation oncologist)
Essentia Health -St. Mary's Medical Center, Duluth, MN	Kenneth J. Dornfeld, M.D. (radiation oncologist)
St. Mary's Medical Center, Huntington, WV	Sanjeev S. Sharma, M.D. (radiation oncologist)
St. Mary's of Michigan, Saginaw, MI	Young H. Kim, M.D. (radiation oncologist)
St. Mary's Regional Medical Center, Reno, NV	Jonathan Tay, M.D. (radiation oncologist/CyberKnife)
Stanford University, Palo Alto, CA	Richard I. Whyte, M.D. (surgeon)
Tianjin Medical University Cancer Institute & Hospital, China	Ping Wang, M.D. (radiation oncologist)
AtlantiCare Regional Medical Center, Egg Harbor Township, NJ	James C. Wurzer, M.D. (radiation oncologist)

Site (ROSEL)	Site Principal Investigator
VU University Medical Center	Suresh Senan F.R.C.R, Ph.D. (radiation oncologist)
University Medical Center Groningen, Groningen, The Netherlands	Harry J.M. Groen M.D., Ph.D. (pulmonologist)
Catharina Hospital, Eindhoven, The Netherlands	Ben van den Borne M.D. (pulmonologist)
Maastricht University Medical Center, The Netherlands	Anne-Marie Dingemans M.D., Ph.D. (pulmonologist)